

**DEVELOPMENT AND OPERATIONS OF THE
ASTROPHYSICS DATA SYSTEM**

NASA Grant NCC5-528

Annual Report

For the Period 1 January 2001 through 31 December 2001

Principal Investigator
Dr. Stephen S. Murray

March 2002

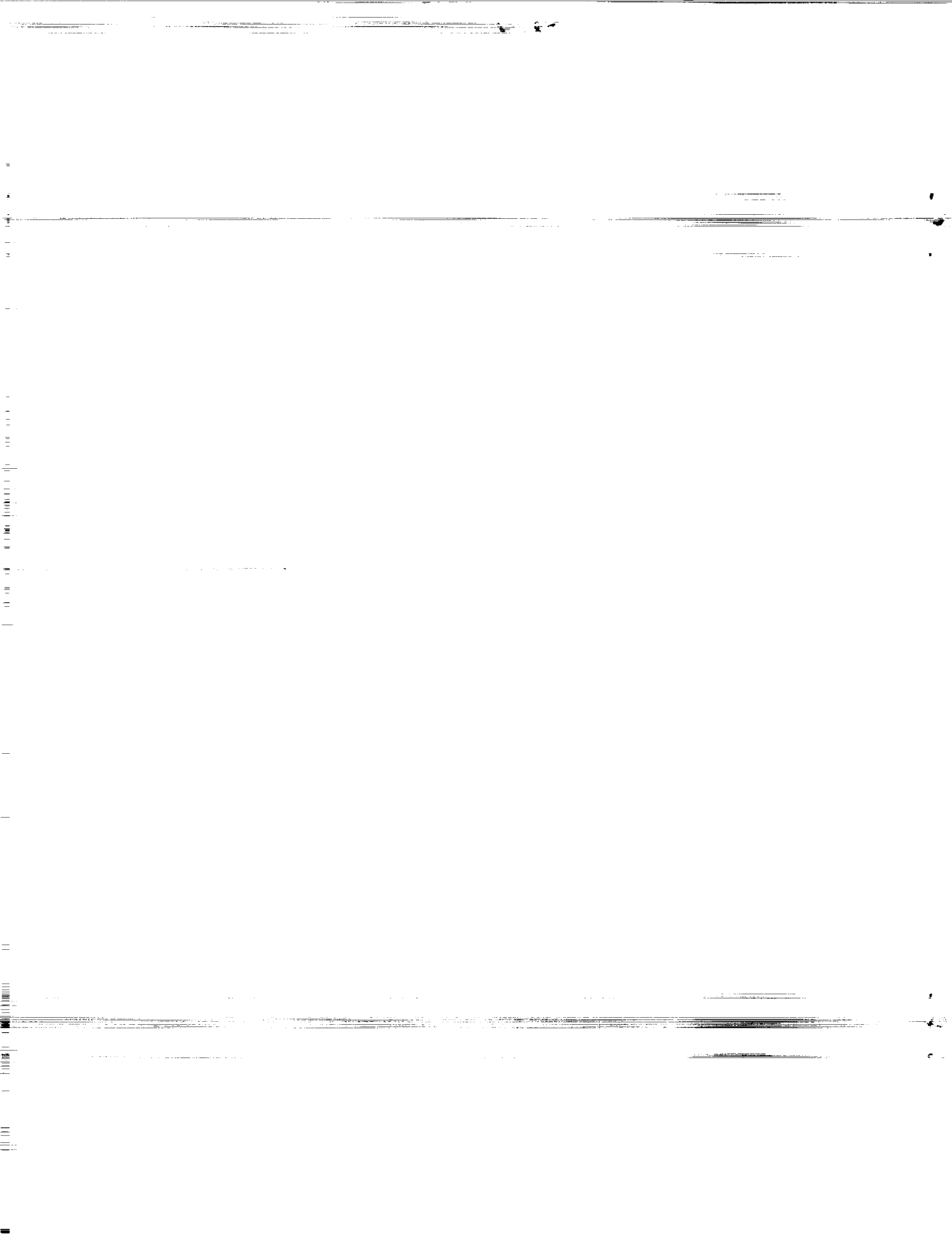
Prepared for:

National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

Smithsonian Institution
Astrophysical Observatory
Cambridge, Massachusetts 02138

The Smithsonian Astrophysical Observatory
is a member of the
Harvard-Smithsonian Center for Astrophysics

The NASA Technical Officer for this grant is Ronald Oliverson, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD 20771



Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 February 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the database, both at SAO and at all mirror sites.
- Coordinated with A&A personnel to fix access to electronic version now at EDP Sciences.
- Resolved problem with thousands of text files in Instrumentation database.
- Corresponded with Science Magazine personnel in order to try to get SGML abstracts from them for inclusion in ADS.

Article Service:

- Transitioned all software used to process scans for the ADS article service to a more efficient storage scheme, and moved the contents of the ADS article archive to the new hierarchy on both main ADS server and mirror servers.
- Modified article processing software to use XML configuration files; implemented publication, volume, and program-specific settings to be used when processing scanned images.
- Integrated the creation of article links in the ADS abstract service by writing scripts that automatically generate lists of scanned pages associated with a particular bibliographic entry.
- Documented changes in the article service interface.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Provided user support to users in the US and abroad.

SUMMARY

The ADS was represented at the AAS meeting in San Diego with an oral talk and a demonstration. The demonstration was hugely successful, especially in light of the fact that the meeting was attended by about 1,200 Physics teachers. This was a group that we had no contact with before. They were very interested in the ADS. Having made contact with this group should help to further propagate knowledge about the ADS in the Physics community and bring it to the attention of young scientists.

Two new mirror sites were established, one in India and in Russia. The site in India is fully operational, the site in Russia is waiting for the data tapes to be loaded. The preprint database is already functional. The mirror site in India is expected to become a full article mirror site soon. The mirror site in Russia is expected to become a partial article mirror site in the near future.

The mirror site in China now has a full article mirror. This is the second full article mirror site after the one in Japan. The mirror site at the CDS in France is working on configuring disks for another full article mirror.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 March 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the database, both at SAO and at all mirror sites.
- Corresponded with UCP personnel about naming conventions for authors with multiple name surnames.
- Corresponded with Kluwer personnel about volumes of Ap&SS which were printed in years other than their stated publication year.

Article Service:

- Updated software used to mirror ADS abstract and article data to the ADS mirror sites to account for changes in the ADS data hierarchy.
- Received and evaluated new DC from Preservation Resources containing scanned microfilm data; processed the scans and made them available for viewing to ADS and library staff.
- Modified Dexter software to account for changes in the ADS article archive.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Configured and populated new hardware to be used for the ADS article and abstract servers, including a new RAID system for scanned articles.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 March 2001

SUMMARY

We installed a new, faster server for the abstract service to improve search speed. The new server should improve query times by about a factor of 2. The article service will move to the previous abstract server next month. This will improve the article service as well.

We started preparations for the Lunar and Planetary Science Conference in March in Houston. We will again have a demo of the ADS there. This year we will also have a poster about the ADS in the science section of the conference.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 April 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Created 4mm DAT tapes containing a full copy of the ADS abstracts and sent it to the new mirror site at INASAN (Russia).
- Modified most ADS's HTML pages to reflect the changes in the server setup.
- Added approximately 15000 MNRAS records to complete database coverage back to volume 1.
- Added over 2000 LPSC meeting and conference abstracts for the upcoming meeting in Houston.
- Corresponded with Elsevier personnel to modify linking syntax and arrange for free display of abstracts for users coming from with the ADS.
- Corresponded with new publishers of PASJ to arrange for data exchange for the ADS.
- Wrote new parser for physics data provided by World Scientific.

Article Service:

- Sent 5 boxes of Astronomische Nachrichten to be scanned.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Presented ADS demonstration at the LPSC conference in Houston.
- Switched the DNS records for the ADS servers to point to the new machines and configured their HTTP and ADS server software.
- Modified most ADS's HTML pages to reflect the changes in the server setup.
- Cleaned up all CGI scripts and ported them to the new servers.

SUMMARY

We were again represented with a demo at the Lunar and Planetary Science Conference in Houston. There was lots of interest in the ADS there. We are starting to see an improvement in the number of people who do not know about it.

This year we had a poster paper about the ADS in the conference. This turned out to be very successful. There was great deal of interest in the poster and we talked with many people about the ADS and made them familiar with the ADS.

The article service was moved to a faster server. This should improve access times. It also reduces the load on the server that handles the SIMBAD US mirror, so access to that service should improve as well.

We set up a mirror of the ADS in Sao Paulo, Brazil. This mirror site will help to server astronomer in South America.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 May 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Wrote automated download procedures for abstract data from Springer and AIP.
- Wrote parser for physics data from Springer, which is automatically downloaded on a weekly basis.
- Summarized current handling of references in the ADS as a starting point towards redesigning reference handling within the ADS.

Article Service:

- Processed scans for all volumes of Solar Physics, Journal of the AAVSO, and the Minor Planet Bulletin.
- Sent 3 boxes of publications to contractor for scanning.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Set up hardware and software for ADS laptop to use a new USB Ethernet interface and an external hard drive for storage of article data, thus making it a fully functional stand-alone ADS mirror site.
- Assisted in introducing the new programmer to the ADS.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 May 2001

SUMMARY

We started to receive scans of the microfilms that were produced by the conservation project at SAO and Harvard. This project is making preservation microfilms of the historical observatory publications. This will allow us to put a set of the astronomical literature on-line that is difficult to get in many libraries.

The CDS in Strasbourg now has a complete article mirror. This will help our European users to better access the scanned articles in the ADS.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 June 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Designed a new processing pipeline used for bibliographic reference parsing, resolution and verification in ADS, which allows the automatic identification of strings containing references to published papers with existing ADS bibliographic records.
- Began populating all reference directories with data provided by the journals and other sources, sometimes modifying original format for consistency.
- Downloaded and added over 2500 abstracts from the spring AGU meeting
- Discussed syntax for adding data catalog descriptions (e.g., ViZier) to the ADS.
- Updated on-line help files on abstract database content for the new ADS Site Map.
- Improved understanding of previous reference harvesting and resolving, including recovering Markus' scripts for references downloading and updating the references with the old harvesting and resolving system.
- Wrote a graphical interface to visualize resolver output

Article Service:

- Updated on-line help files on article printing.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Prepared demo on ADS laptop for metadata entry on scans digitized from the Wolbach Library preservation microfilms.
- Registered DNS record for host ads.si.edu to point to the main ADS site.
- Prepared handout for the AGU meeting in Boston.
- Attended the AGU spring meeting in Boston and gave demonstrations of the ADS services at the exhibit center.
- Installed latest version of Python and setup of Python so that scripts can run seamlessly from Linux or Solaris platforms.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray

Status as of: 1 June 2001

Achievement: _____ G. Eichhorn (SAO)

SUMMARY

We started to receive more Physics abstracts from different publishers (e.g., Springer, World Scientific). We are in discussions with the APS to get their abstracts for PhysRev. This will greatly enhance the ADS Physics database.

We started preparations for the AAS meeting (a poster paper), the AGU meeting (a demo), and the SLA meeting (a talk). In addition, the ADS was invited to present a talk at the United Nations Workshop for Basic Space Science in the Developing World in June.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 July 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Established the technical details for the periodic downloads of abstract and reference data from the American Physical Society (APS); proposed agreements on data-sharing policies; oversaw implementation of software allowing downloads based on the Open Archives Initiative Metadata Harvesting Protocol.
- Isolated problem with chemical symbol capitalization and modified text files in the ADS accordingly to improve searching.
- Reorganized reference sources under /proj/ads/references.
- Started writing python classes for an automatic and configurable reference handler.
- Included back issues of Science and linked all articles on-line at their web site.

Article Service:

- Implemented new procedures to manage embargo of full-text articles in the ADS article service as per the agreements negotiated with the copyright holders.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Prepared handouts for AAS and SLA meetings.
- Prepared an invited paper to be presented at the bi-annual Human Computer Interaction conference.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 July 2001

SUMMARY

We presented a poster paper at the AAS meeting, a talk at the SLA meeting, and a talk at the United Nations workshop. Guenther Eichhorn received an award at the SLA meeting for work on the ADS. We also had a demo at the AGU meeting. We will continue talking with the AGU in order to get the JGR abstracts.

Work on a program to capture metadata for the scans of historical observatory publications from preservation microfilms started. We will need help from outside users with this project. The microfilm scans do not have any information about page numbers of articles associated with them. We will have a graphical user interface that will allow interested users to help us with the metadata capture.

We started to rework our help pages. This will make it much easier to find relevant help information in the ADS.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 August 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Incorporated 330,000 references into the physics database from the Physical Review Series.
- Incorporated new dataset of approximately 12,000 references on helioseismic articles from GONG.
- Fixed capitalization problem and improved case-sensitive searching.
- Continued gathering of new material and shipments of table of contents to our contractor in India for typing in and including in the ADS.
- Modified syntax of tagged format to allow for greater flexibility in keyword sets and language tags.
- Set up and populated the new ADS mirror in India (at IUCAA).
- Implemented creation of ToC links in an automated fashion from a list of conference proceedings bibstems.
- Implemented new mail filters for the email addressed to ADS in order to allow automatic logging of output from procedures run on mirror sites and bounces from the ADS article email retrieval interface.
- Assisted in implementing the parsing pipeline for the APS data.
- Worked on new help_pages and site_map documentation, including creating new scripts to convert the text files into HTML pages.

Article Service:

- Finished a first version of the Historical metadata interface.
- Finished a first running version of the automatic reference handler.

Preprint Service:

- Updated software used to rasterize and index the CfA preprints to account for changes in the ADS abstract and article service software.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 August 2001

SUMMARY

We have an agreement with the APS to get their abstracts and reference lists for all articles in PhysRev back to volume 1. We started downloading of the abstracts. This will greatly increase the coverage of the ADS Physics database.

Our mirror sites in India and China now have a full article mirror.

We started preparations for attendance at the ADASS meeting with one talk and one poster paper.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 September 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Incorporated over 7500 new references from ICR conference series and performed many corrections to existing data in the database.
- Incorporated two sets of PhdD theses into the astronomy database.
- Collaborated with personnel at Nature, PASP, JBAA, STScI, and ARI to improve flow of new data and linking of old data.
- Streamlined the creation of links from ADS bibliographic records to internal and external resources by providing pre- and post-processing capabilities for each type of object being linked.
- Provided tools and documentation allowing the seamless integration of ADS-generated BibTeX records into modern authoring packages.
- Created new classes and scripts for downloading references and abstracts from APS.
- Wrote a dedicated reference resolver for digesting the large volume of data from APS.

Article Service:

- Processed early volumes of MNRAS and made the scans available on-line.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Attended the conference Human Computer Interaction 2001 and presented the invited paper "Information Agents in Astronomy".
- Tested public domain software for on-line incremental backups.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 September 2001

SUMMARY

Downloading of the APS data continued. We started processing of their reference lists. This will add a lot of references to our citation database. Preparations for the ADASS conference continued.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 October 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Incorporated thousands of book reviews from the University of Toronto Library, including matching them to similar records in the database, revising author handling to include book authors and review authors, and correcting mistakes in existing book review entries.
- Reviewed, expanded and updated help pages describing the use of the ADS article service and the ADS preference setting system.
- Reviewed, expanded and updated help pages describing the data and results aspects of the ADS.
- Created bibliographic code abbreviations for first set of microfilms being included in the ADS.
- Updated procedures that create inks to APS abstracts, references, and full-text articles.

Article Service:

- Processed additional CDs created from microfilm scanning.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 October 2001

SUMMARY

We attended the ADASS meeting with one talk and one poster paper. Preparations for demos at the DPS meeting and the AGU meeting started. We will also have a poster paper at the DPS meeting. Members of the planetary science community are still not as familiar with the ADS as members of the astrophysics community. Attendance at the DPS meetings will hopefully improve this situation.

We finished downloading the APS data. This added 300,000 entries to our Physics database. Their reference lists added 1.5 million reference lists to the ADS citations database.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 November 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Added over 28000 BAAS abstracts to the ADS, including matching of several years of electronic AAS abstracts with their printed counterparts.
- Revised naming convention for Vizier catalog descriptions in response to user feedback.
- Requested, received and incorporated several years of back issues for Elsevier journals, in particular to improve coverage of their physics journals.
- Updated and extended tools to facilitate linking and integration of ADS services, including a PERL library and sample scripts for bibcode verification, reference resolution, and abstract formatting, the software is currently used by several publishers to create links to ADS.
- Continued working on APS parsing, and finished the complete download of the APS dataset.
- Implemented several improvements to the general reference handler classes.
- Started writing an automatic abstract downloader.
- Wrote resolveserver, an xmlrpc server to access the reference resolver fro a cgi-bin interface with a reasonable handling of the computer resources.

Article Service:

- Processed over 50 scanned volumes of IAU Symposia, and made full-text available on-line.
- Retrieved old publications to be scanned from the warehouse.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Attempted registration of astrophysics.info and astrophysics.org (without success).
- Attended ADASS XI and wrote article for proceedings.

SUMMARY

Preparations for the AAS meeting in January started. We will have an invited talk in the session for historical astronomy, a demo and poster paper. The talk will present the efforts for making the historical observatory literature available on-line and will show a working version of the metadata capture program.

We put more scans of historical observatory publications scanned from microfilms on-line. Work on the metadata capture program continued.

The re-built Help pages are now on-line. They make it much easier to find help with different features of the ADS.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 December 2001

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Established policy on author-submitted citations.
- Modified mirroring software in an attempt to provide fine-tuned options allowing efficient replication of individual files.
- Finished the abstract downloader.
- Integrated all existing Journal downloading into the system.
- Wrote classes to handle downloading status and configuration files.
- Worked with Wiki: general StructuredText python classes to convert help text files into documentation, HTML and PDF/PS formats.

Article Service:

- Collaborated with staff from the Wolbach library at the CfA in trying to complete microfilming of historical literature not locally available.
- Processed scans for 6 new journals, including Zeitschrift fur Astrophysik, The Observatory, and Antarctic Meteorite Research.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Upgraded ADS laptop PC to the latest version of RedHat, and configured it to run as an ADS mirror.
- Worked with new staff member to assign tasks, and summarize status of existing tasks.
- Prepared handout for and demonstrated the ADS at the DPS meeting in New Orleans.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 December 2001

SUMMARY

We completed preparations for our poster paper and demo at the DPS meeting. The DPS meeting is in late November/early December. Preparations for the AGU meeting continued. Contacts with the AGU indicate that we may get their data soon. Preparations for the AAS meeting continued. In particular, the work on the metadata capture program continued so a working version will be available during the talk at the Historical Astronomy Division.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 January 2002

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Added links to ScienceDirect as harvested from published table of contents.
- Added links to all BAAS full-text papers requiring significant corrections to existing data.
- Investigated data harvesting methodologies for downloading data from Library of Congress and the ArXiv preprint archive.
- Implemented a flexible program to generate statistics about ADS data holdings.
- Assisted staff at ESO, CDS and BAO in transitioning the ADS mirror machine to new servers.
- Added automated procedure for downloading Springer physics books.
- Adapted Wiki to generate website pages and mirrors.

Article Service:

- Created a suite of new procedures used to integrate scanned data from the historical literature microfilms and processed over 200000 pages.
- Worked on a major rewrite of the historical metadata interface.

Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Prepared handout for and demonstrated the ADS at the AGU meeting in San Francisco
- Updated tucana to red-hat 7.2 after a system failure.

SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 January 2002

SUMMARY

We attended the DPS meeting (demo and poster paper) and the AGU meeting (demo). There was lots of interest in the ADS at the AGU meeting. We reached an agreement with the AGU to receive their abstracts starting with the January 2000 abstracts in February 2002.

Preparations for the AAS meeting were completed. The metadata capture program will be ready for demonstration at the AAS meeting.

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 February 2002

DEVELOPMENT

SAO

TASKS ACCOMPLISHED:

Abstract Service:

- Continued regular updates of abstracts in the databases, both at SAO and at all mirror sites.
- Established a new naming convention of QB books in preparation for adding physics books from Hollis or Library of Congress.
- Modified handling of object tag so as not to interfere with XHTML definition.
- Worked on moving "what's new" announcements to a majordomo email list so as not to interfere with divisional mail handling.
- Implemented and tested new first author feature following suggestions from users at the AAS meeting.
- Added SSRv entries back to volume 1 in preparation for scanning of the journal.
- Assisted in the re-configuration of the ADS mirror site at the CDS and sent a new set of tapes containing article data to allow re-creation of the ADS article data lost during the move.
- Created scripts to automatically download Astrobiology.

Article Service:

- Made available on-line links to all 30 issues of the Bulletin of the American Astrophysical Society scanned in the ADS article service.
- Sent 5 boxes of publications to be digitized to scanning company.
- Continued working on metadata interface and scripts for handling user-inputted metadata, and post-processing of metadata.

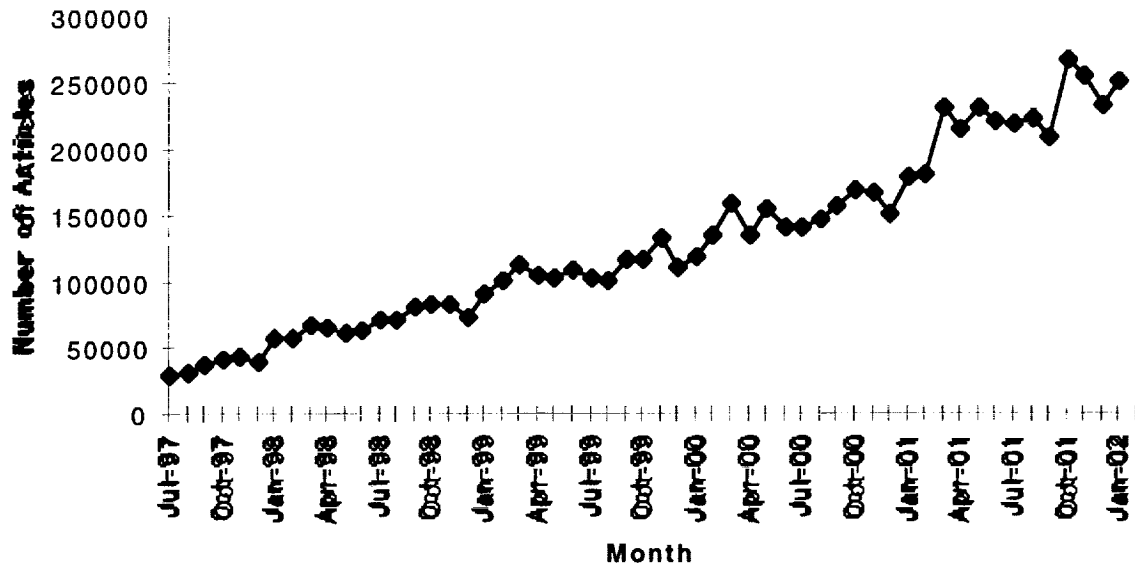
Miscellaneous:

- Responded to user comments about the Abstract and Article Services.
- Prepared handout for and demonstrated the ADS at the AAS meeting in Washington, DC.
- Reconfigured data holdings on older ADS server. Planned move of disk and tape drives to new hardware.
- Spent time on system administration on tucana following a disk failure.

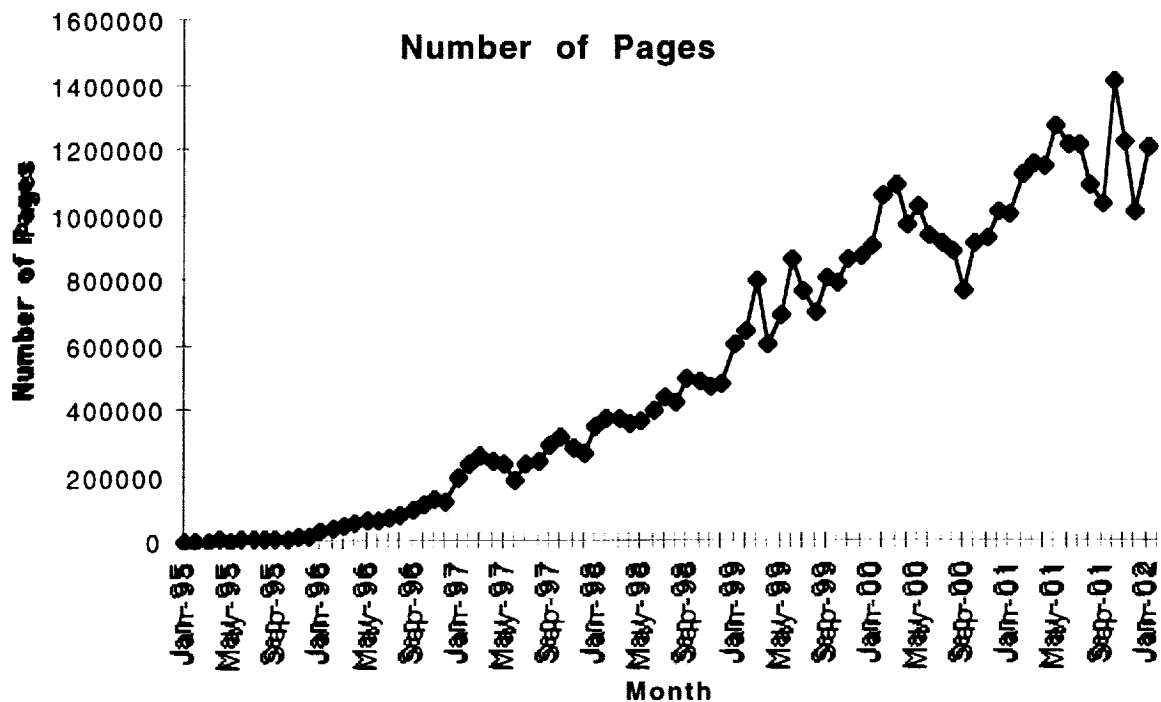
Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 February, 2002

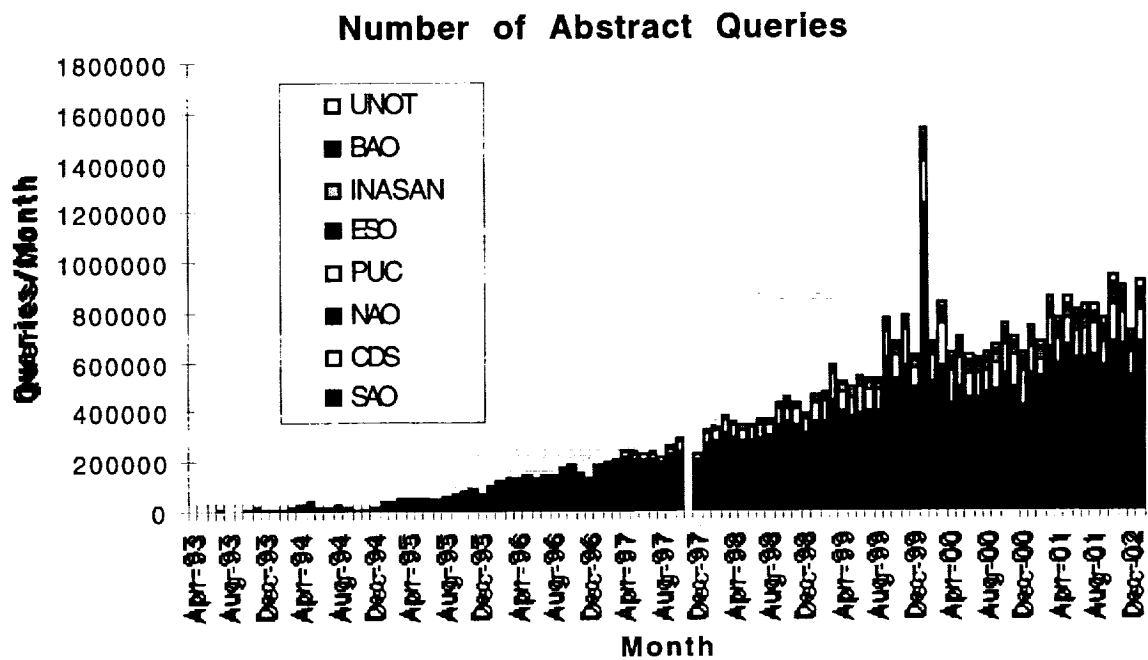
Number of Articles



Number of Pages



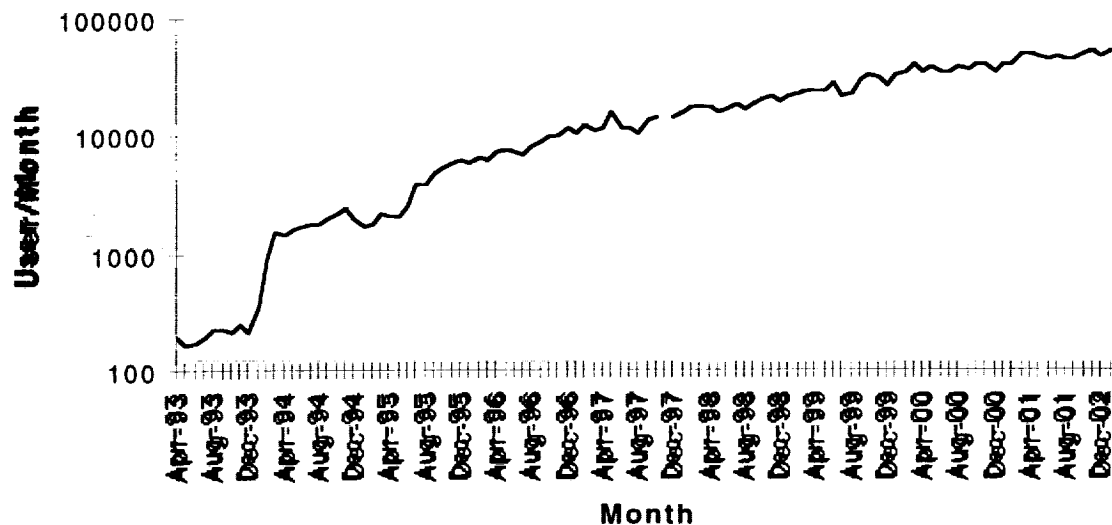
Status as of: 1 February, 2002



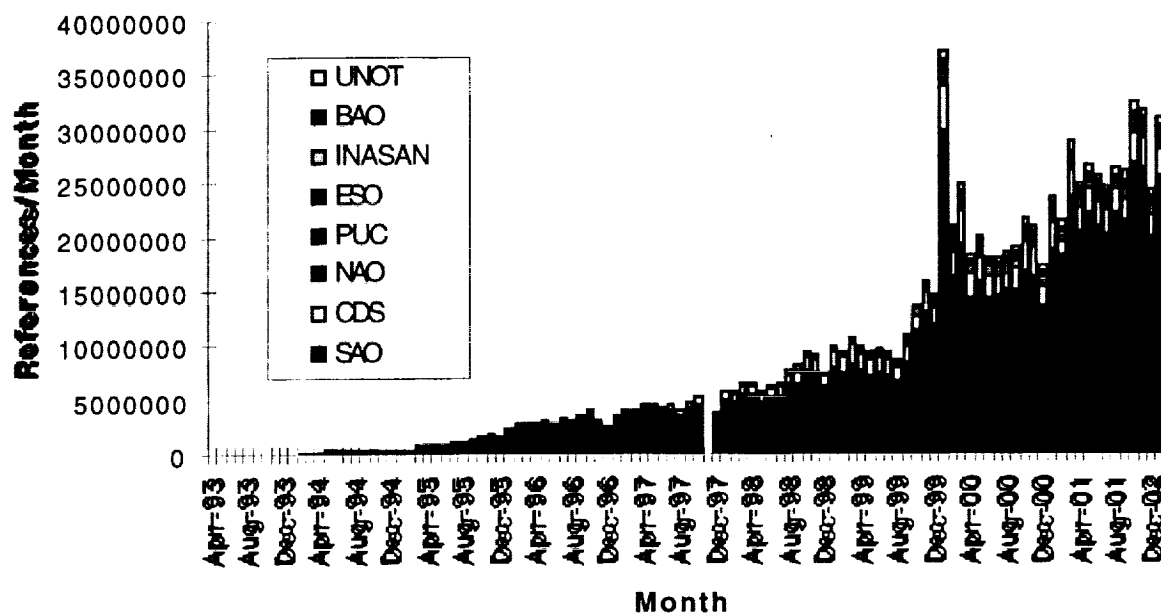
Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 February, 2002

Distinct Users, Abstract Service



Number of Retrieved References



SAO

ASTROPHYSICS DATA SYSTEM

Approved: _____ S. Murray
Achievement: _____ G. Eichhorn (SAO)

Status as of: 1 February 2002

SUMMARY

We attended the AAS meeting with an invited talk, a poster paper, and a demo. It was a very successful meeting. The demonstration of the metadata capture program has generated interest in helping up with this project. We have already several people working with the program and sending us metadata. This will make the historical literature collection much more useful for our users.

